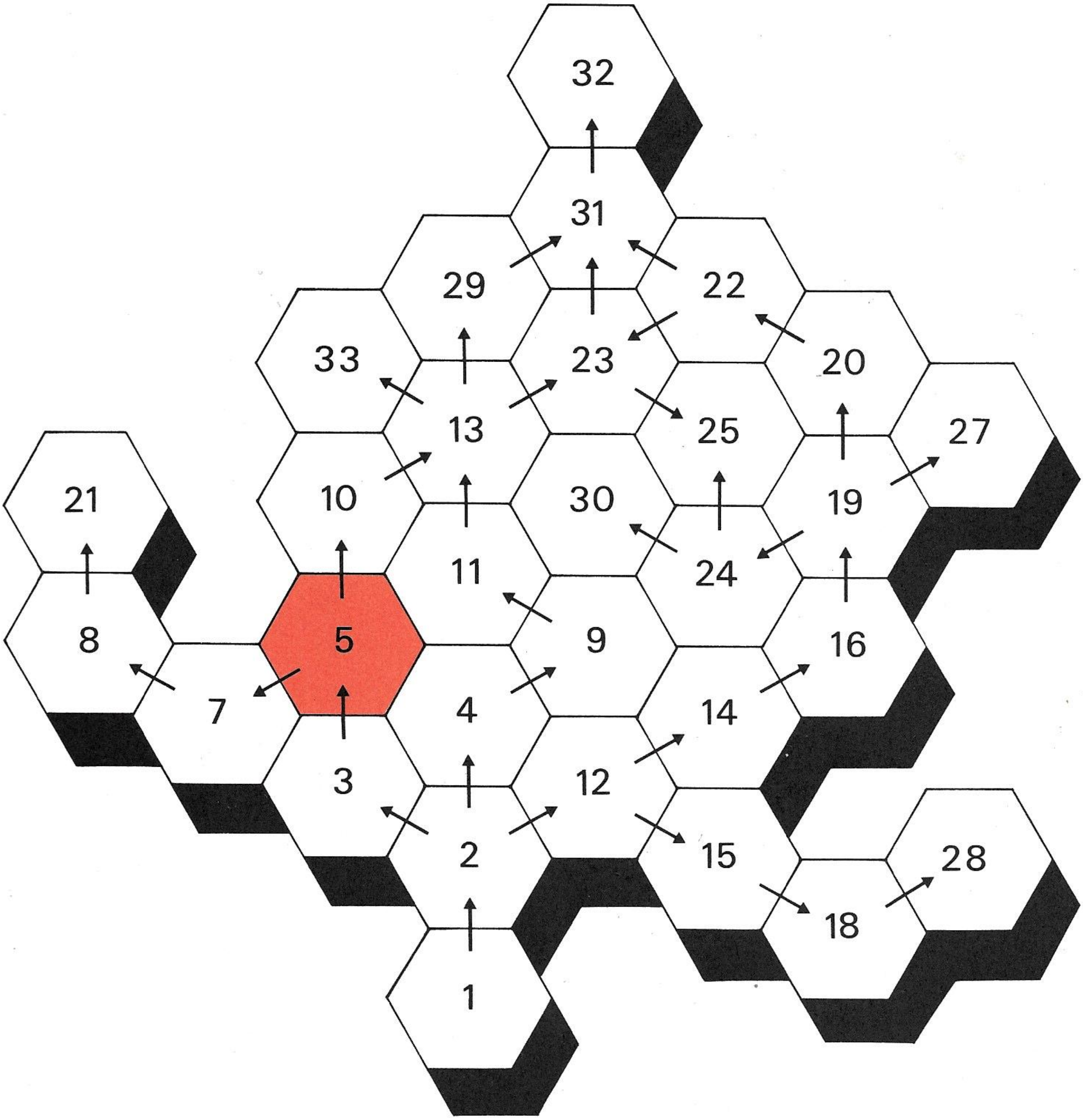
M201 5

THE OPEN UNIVERSITY

Mathematics: A Second Level Course

Linear Mathematics Unit 5

Determinants and Eigenvalues



The Open University

Mathematics: A Second Level Course

Linear Mathematics Unit 5

DETERMINANTS AND EIGENVALUES

Prepared by the Linear Mathematics Course Team

The Open University Press

The Open University Press
Walton Hall, Milton Keynes.
MK7 6AA

First published 1972. Reprinted 1976

Designed by the Media Development Group of the Open University.

Printed in Great Britain by
Martin Cadbury

SBN 335 01094 6

This text forms part of the correspondence element of an Open University Second Level Course. The complete list of units in the course is given at the end of this text.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, P.O. Box 81, Walton Hall, Milton Keynes, MK7 6AT.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Walton Hall, Milton Keynes, MK7 6AB.

1.2

Contents

Set Books

D. L. Kreider, R. G. Kuller, D. R. Ostberg and F. W. Perkins, *An Introduction to Linear Analysis* (Addison-Wesley, 1966).

E. D. Nering, *Linear Algebra and Matrix Theory* (John Wiley, 1970).

It is essential to have these books; the course is based on them and will not make sense without them.

Conventions

Before working through this correspondence text make sure you have read *A Guide to the Linear Mathematics Course*. Of the typographical conventions given in the Guide the following are the most important.

The set books are referred to as:

K for *An Introduction to Linear Analysis*
N for *Linear Algebra and Matrix Theory*

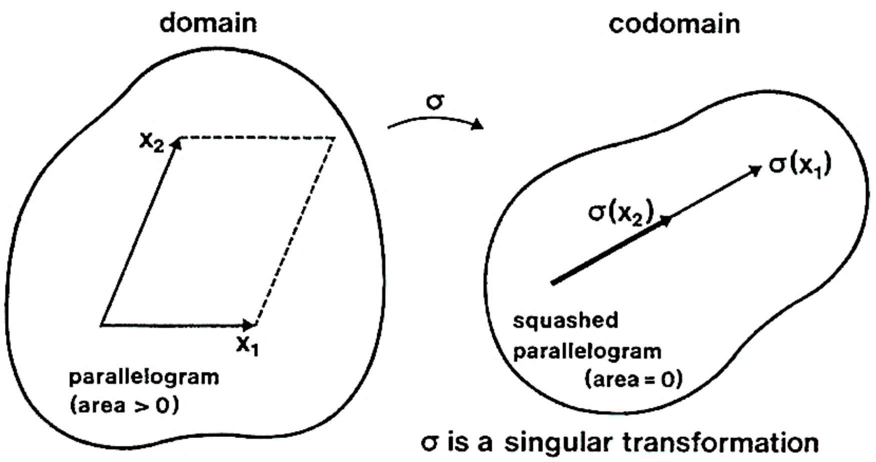
All starred items in the summaries are examinable.

References to the Open University Mathematics Foundation Course Units (The Open University Press, 1971) take the form *Unit M100 3, Operations and Morphisms.*

5.0 INTRODUCTION

In this unit we look at linear transformations whose domain and codomain are the same vector space; such linear transformations we call *endomorphisms*. We shall be looking for properties that are "geometrical" (i.e. intrinsic) in character, in the sense that they do not depend on the basis used in representing the transformation as a matrix. We have already seen some "geometrical" properties of this kind—in particular, the dimension of the domain vector space of the transformation and the rank (the dimension of the image space) of the transformation.

We shall consider first a number which measures the factor by which n-dimensional "volumes" in a real vector space are changed by a linear transformation. This factor is called the *determinant* of the transformation; the immediate use we shall be making of it is to give a convenient *theoretical* criterion for identifying singular transformations (transformations which do not have inverses). It turns out that such transformations are characterized by having determinant zero. The reason is that a singular transformation maps its domain on to a subspace of lower dimension; the effect of such a transformation with a two-dimensional domain is to convert a plane figure into a line, so reducing the "volume" (area) of the plane figure to zero.



σ is a singular transformation

The second type of number we shall consider measures the effect of a linear transformation on lengths, rather than volumes. Although the idea of length is not part of the definition of a vector space, it is possible to compare the lengths of two vectors *provided the vectors are scalar multiples of one another*. Geometrically, this means that the two vectors must be "parallel." Thus if the linear transformation L maps a particular vector $\mathbf{x}$ to a vector $\lambda\mathbf{x}$, where λ is a scalar, then we may say that it has changed the "length" of this vector by a factor λ. The special vectors that have this property are called *eigenvectors* of L and the corresponding scalar multipliers are called *eigenvalues* of L. The one-dimensional subspace spanned by an eigenvector is invariant under L; i.e. it maps to itself. It is very often the case in mathematics that transformations are best studied by observing and making use of entities which they leave unchanged: that proves to be the case here, as we shall see.

One of the main applications of eigenvalues and eigenvectors is in the theory of vibrating mechanical or electrical systems, such as the aircraft vibrations considered in the radio programme of *Unit M100 31, Differential Equations II.* We shall be studying this type of application later in the course.

Another application arises in quantum physics; for example, the reason for the ghastly yellow colour produced by sodium-vapour street lamps is

that sodium emits visible light of only one frequency. The theory of this kind of light emission relates the possible frequencies of emitted light to the possible values for the energy of a sodium atom, which are given by the eigenvalues of a certain linear operator. (This course does not, however, require any knowledge of quantum mechanics.) There are many other applications, some of which you will meet later in the course.

In this unit we shall be looking mainly at the theory, rather than the applications of eigenvalues. We shall introduce the determinant to provide a criterion for singular transformations; this leads to a simple method for finding the eigenvalues and eigenvectors in spaces of dimension 2 or 3. In later units we shall find many applications for eigenvalues and eigenvectors, and shall look at effective methods for calculating them in spaces of many dimensions.

5.1 THE DETERMINANT—AN INTRINSIC PROPERTY OF ENDOMORPHISMS

5.1.0 Introduction

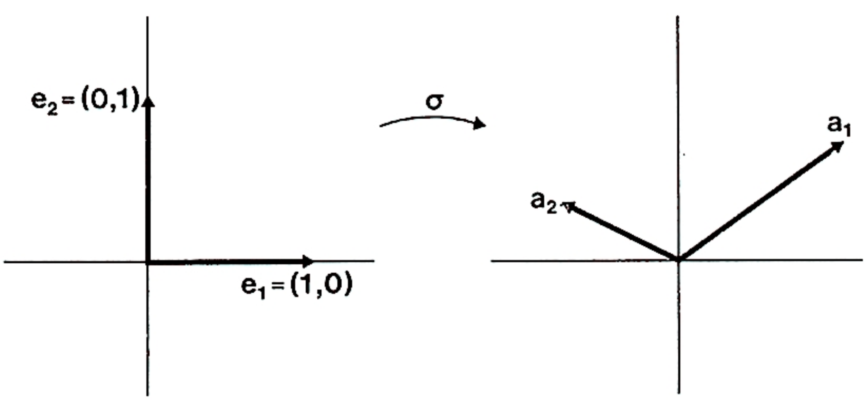
In this section of the unit we shall define, for any endomorphism of an n-dimensional vector space, a scalar called the determinant of that endomorphism. It can be interpreted as the factor by which the endomorphism magnifies n-dimensional "volumes". We shall approach the definition by considering endomorphisms of a plane into itself, for which case these "volumes" are really areas, and look for properties of the magnification factor that can be generalized to n-dimensional vector spaces. These properties are then formalized as a set of axioms which are the basis of the definition of determinants. We then obtain an algorithm for evaluating determinants, which makes use of what you already know about Hermite normal form.

5.1.1 The Magnification of Areas

The two-dimensional example we shall consider is the vector space R^2, where elements can be represented in the usual way by points in a plane. We saw in *Unit 2, Linear Transformations* (*Theorem 1.17*, page N34) that any linear transformation is completely specified by its effect on a fixed basis, and so any endomorphism of the plane can be specified by giving the image vectors of some fixed basis of R^2. We shall take as fixed basis, the standard basis

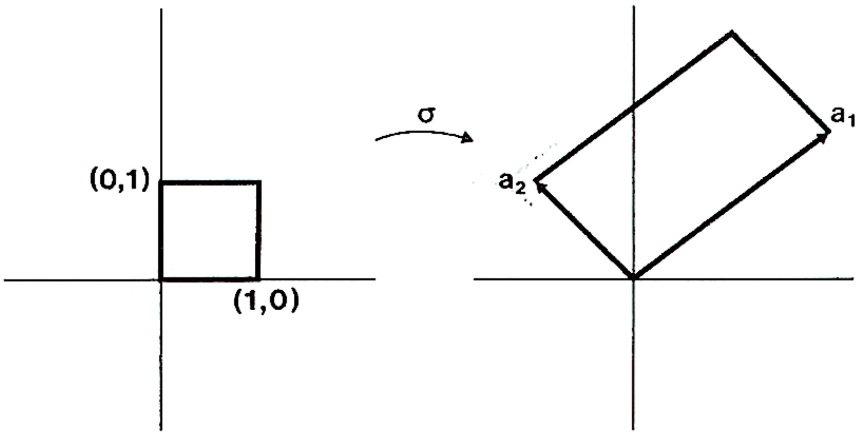
$$\{(1, 0), (0, 1)\}.$$

To conform with the notation of **K** we shall denote this basis by $\{\mathbf{e}_1, \mathbf{e}_2\}$, and its image by $\{\mathbf{a}_1, \mathbf{a}_2\}$.



To calculate the magnification produced by the endomorphism σ, we compare the area of a geometrical figure in R^2 with the area of its image.

A convenient figure to consider is the square shown in the diagram. The image of this square under σ is the parallelogram shown.



If we take the square to be of unit area, then the area of the parallelogram is equal to the magnification produced by the linear transformation. It is not too difficult to obtain a formula for the area of the parallelogram in terms of the coordinates of $\mathbf{a}_1$ and $\mathbf{a}_2$, but the generalization of this formula to n dimensions is not obvious. In keeping with our philosophy of looking for basis-independent properties, we can approach the problem from the point of view of vectors rather than coordinates.

We are interested in the function

$$D\colon (\mathbf{a}_1, \mathbf{a}_2) \longmapsto \text{area of parallelogram} \quad (\mathbf{a}_1, \mathbf{a}_2 \in R^2).$$

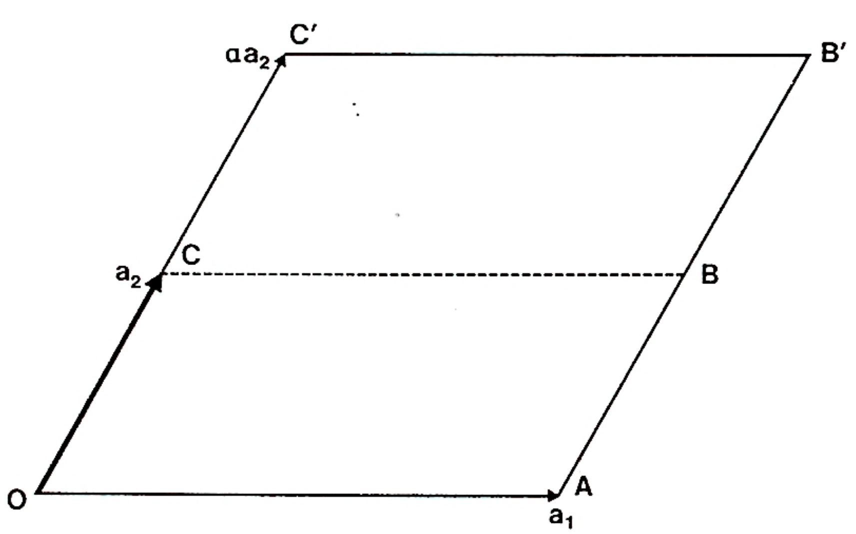
(We use the letter D because we are using this area as our definition of *determinant* for endomorphisms of R^2. There should be no confusion with the use of the letter D for the differentiation operator.)

Since the domain of the function D is a set of *pairs* of vectors, we begin by simplifying and keep one of the vectors fixed and see how the area depends on the other; in other words, we study the function

$$f\colon \mathbf{a}_2 \longmapsto D(\mathbf{a}_1, \mathbf{a}_2) \qquad (\mathbf{a}_2 \in R^2)$$

for some fixed vector $\mathbf{a}_1$. Let us investigate whether this function is a linear transformation.

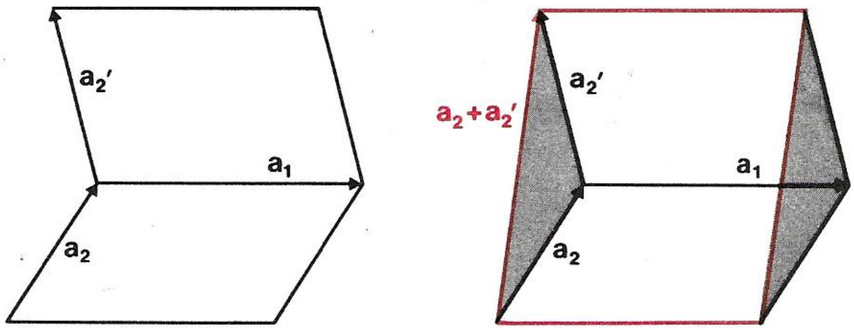
To test whether the function f preserves scalar multiplication, we compare two parallelograms on the same base $\mathbf{a}_1$ with neighbouring sides $\mathbf{a}_2$ and $\alpha\mathbf{a}_2$ respectively, where α is a scalar.



From elementry geometry we know that the parallelogram $AB'C'O$ has area α times that of $ABCO$, so that we have

$$f(\alpha \mathbf{a}_2) = \alpha f(\mathbf{a}_2).$$

To test whether f is a morphism for addition, we can put together two parallelograms on the same base as shown below:



Since the two shaded areas are equal, the two parallelograms with black sides have the same total area as the parallelogram with red sides; that is,

$$f(\mathbf{a}_2) + f(\mathbf{a}_2') = f(\mathbf{a}_2 + \mathbf{a}_2').$$

Thus we have shown that f is a linear transformation. In the same way we can show that the function

$$\mathbf{a}_1 \longmapsto D(\mathbf{a}_1, \mathbf{a}_2) \qquad (\mathbf{a}_1 \in R^2),$$

where $\mathbf{a}_2$ is any fixed vector in R^2, is a linear transformation. Thus $D(\mathbf{a}_1, \mathbf{a}_2)$ is linear in $\mathbf{a}_1$ for fixed $\mathbf{a}_2$, and in $\mathbf{a}_2$ for fixed $\mathbf{a}_1$. (Such a function is said to be *bilinear*; we shall study bilinear functions in more detail later in the course.)

There are two more properties of the area magnification function which are geometrically obvious. First, if $\mathbf{a}_1 = \mathbf{a}_2$, then the parallelogram collapses and has zero area:

$$D(\mathbf{a}_1, \mathbf{a}_1) = 0 \qquad (\mathbf{a}_1 \in R^2).$$

Secondly, if $\mathbf{a}_1$ and $\mathbf{a}_2$ are the same as our original basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, then our transformation must be the identity transformation, which leaves areas unchanged, and therefore

$$D(\mathbf{e}_1, \mathbf{e}_2) = 1.$$

It is possible to show that there is precisely one function D with the three properties we have discussed, *viz*:

(i) D is bilinear
(ii) $D(\mathbf{a}_1, \mathbf{a}_1) = 0$ $(\mathbf{a}_1 \in R^2)$
(iii) $D(\mathbf{e}_1, \mathbf{e}_2) = 1$

(The proof is given in a reading passage which we shall refer to shortly.) This function is specified by

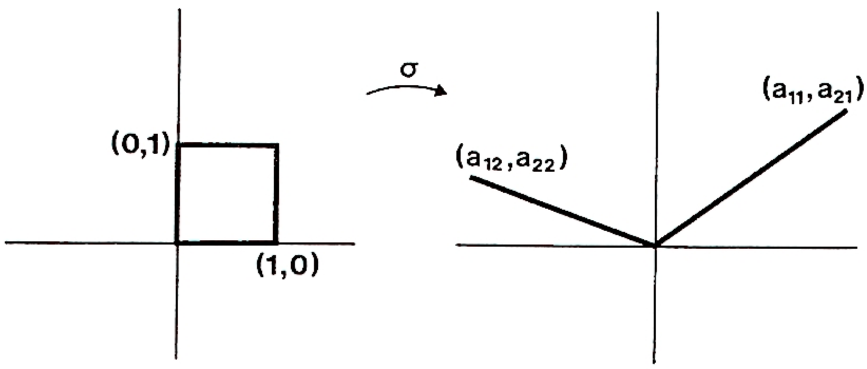
$$D(\mathbf{a}_1, \mathbf{a}_2) = a_{11}a_{22} - a_{12}a_{21}$$

where a_{11} and a_{21} are the coordinates of $\mathbf{a}_1$, and a_{12} and a_{22} are the coordinates of $\mathbf{a}_2$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$. Note that D is a transformation from R^2 to R.

Since $\mathbf{a}_1$ and $\mathbf{a}_2$ are the images of $\mathbf{e}_1$ and $\mathbf{e}_2$ under the endomorphism σ, we can represent σ, relative to $\{\mathbf{e}_1, \mathbf{e}_2\}$, by the matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

This is because the coordinates of the images of the two basis vectors in the domain form the two columns of the matrix of σ (see sub-section 2.2.1 of *Unit 2, Linear Transformations*).



Exercise

From geometrical considerations, what would you expect the determinants of the following endomorphisms of R^2 to be?

(i) A rotation through the angle θ radians about the origin.
(ii) A magnification of all linear dimensions by 2.
(iii) A reflection in the y-axis.

Obtain the matrices of these transformations (by considering the images of the basis vectors, as in *Unit 2, Linear Transformations*), calculate their determinants from the formula given above, and so check your geometrical intuition.

Solution

(i) The matrix is $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ as given in *Unit 2, Linear Transformations*. Its determinant is $\cos^2\theta + \sin^2\theta = 1$, confirming that rotations do not alter areas.

(ii) The matrix is $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, and its determinant is 4 (not 2): doubling the linear dimensions quadruples the area.

(iii) The matrix is $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, and its determinant is -1, not $+1$ as you may have expected. Although the transformation leaves the areas of figures unaltered, it does "turn them over", just as a mirror interchanges right and left. This is the significance of the negative sign.

5.1.2 Defining the Determinant Function

In this sub-section we see how to generalize the idea of determinant to endomorphisms in the vector space R^n. The method is to generalize the three properties we found in sub-section 5.1.1 for area magnification factors in R^2, and use these three as axioms for a definition of determinants in general.

READ Section III-1 on pages K680–681.

Notes

(i) *lines 6 to 7, page* K680 Strictly speaking, D (not $D(\mathbf{a}_1, \ldots, \mathbf{a}_n)$) is the function; its codomain is $\mathscr{R}$ (this is the significance of "real-valued") and its domain is the set of all n-tuples of vectors from $\mathscr{R}^n$. Although the discussion in K is in terms of real vector spaces, the corresponding theory for vector spaces over an arbitrary field (with $D: F^n \times \ldots \times F^n \longmapsto F$) is analogous.
(ii) *line 8, page* K680 This is a generalization of the condition $D(\mathbf{a}_1, \mathbf{a}_1) = 0$ we found in the preceding section; it makes the determinant function a useful theoretical criterion for linear independence of vectors.
(iii) **Definition III-1**, *page* K680 The three conditions in this definition are generalizations of the three properties of $D(\mathbf{a}_1, \mathbf{a}_2)$ we found previously for $\mathscr{R}^2$. Note that we have not yet shown whether a function satisfying these conditions exists, nor whether it is unique if it does exist, nor how to calculate it if it does exist and is unique.
(iv) *line 10, page* K681 "Since the function D..." In effect, this paragraph defines a new function, whose domain is the set of all $n \times n$ square matrices with real entries, whereas the domain of D is the set of all n-tuples of vectors in $\mathscr{R}^n$.

The relation between the functions det and D is indicated by the following example:

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = D((1, 3), (2, 4))$$

The image of a matrix M under the function det is called the *determinant* of M. Note that we have only defined the determinant of a matrix M with real entries.

Exercises

1. Evaluate $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix}$, i.e. $\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

   with the help of Conditions I and III in **Definition III-1.**

2. Use the preceding result, and Conditions I and II, to evaluate

$$\begin{vmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix}$$

3. Use Conditions I, II and III to evaluate

$$\begin{vmatrix} a & b & c \\ 0 & d & f \\ 0 & 0 & g \end{vmatrix}$$

   where all the letters stand for real numbers.

Solutions

1. By the definitions of det and of the standard basis vectors in R^3, we have

$$\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = D(\mathbf{e}_1, 2\mathbf{e}_2, 3\mathbf{e}_3)$$

$$= 2D(\mathbf{e}_1, \mathbf{e}_2, 3\mathbf{e}_3) \quad \text{(Condition I)}$$

$$= 6D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \quad \text{(Condition I)}$$

$$= 6 \quad \text{(Condition III).}$$

2. $$\begin{vmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} = D(\mathbf{e}_1, 2\mathbf{e}_2, 4\mathbf{e}_1 + 3\mathbf{e}_3)$$

$$= D(\mathbf{e}_1, 2\mathbf{e}_2, 4\mathbf{e}_1) + D(\mathbf{e}_1, 2\mathbf{e}_2, 3\mathbf{e}_3)$$
(Condition I)

$$= 4D(\mathbf{e}_1, 2\mathbf{e}_2, \mathbf{e}_1) + D(\mathbf{e}_1, 2\mathbf{e}_2, 3\mathbf{e}_3)$$

$$= 0 + D(\mathbf{e}_1, 2\mathbf{e}_2, 3\mathbf{e}_3)$$
(Condition II)

$$= 0 + 6$$
(Solution 1)

$$= 6.$$

3. $$\begin{vmatrix} a & b & c \\ 0 & d & f \\ 0 & 0 & g \end{vmatrix} = D(a\mathbf{e}_1, b\mathbf{e}_1 + d\mathbf{e}_2, c\mathbf{e}_1 + f\mathbf{e}_2 + g\mathbf{e}_3)$$

$$= a\,D(\mathbf{e}_1, b\mathbf{e}_1 + d\mathbf{e}_2, c\mathbf{e}_1 + f\mathbf{e}_2 + g\mathbf{e}_3)$$
(Condition I)

$$= a[bD(\mathbf{e}_1, \mathbf{e}_1, c\mathbf{e}_1 + f\mathbf{e}_2 + g\mathbf{e}_3) + d\,D(\mathbf{e}_1, \mathbf{e}_2, c\mathbf{e}_1 + f\mathbf{e}_2 + g\mathbf{e}_3)]$$
(Condition I)

$$= adD(\mathbf{e}_1, \mathbf{e}_2, c\mathbf{e}_1 + f\mathbf{e}_2 + g\mathbf{e}_3)$$
(Condition II)

$$= ad[cD(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1) + fD(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2) + gD(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)] \quad \text{(Condition I)}$$

$$= adg\,D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \quad \text{(Condition II)}$$

$$= adg \quad \text{(Condition III).}$$

Thus the determinant of an upper triangular matrix (one having only zeros below its main diagonal from top left to bottom right) equals the product of the diagonal elements. The same result holds for lower triangular matrices.

5.1.3 Column Operations

The definition we have adopted for the determinant function leaves several questions unanswered. Does such a function exist? If so, is it unique? If it does exist, how do we calculate images under the function? The first two questions are answered in **Theorem III-3** on page **K684**; we shall ask you to read this theorem and its proof, but we shall not expect you to reproduce it; all you need for this course is the fact that the determinant function does exist, some idea of its form, and that it is unique. The third question, on the other hand, is much more important: in order to be able to use determinants, we shall need to be able to evaluate them.

The basis of the *practical* method for evaluating determinants is the same as that of the method you studied in *Unit 3, Hermite Normal Form* for calculating inverses of matrices; it is the fact that any non-singular matrix can be expressed as a product of elementary matrices. Since a determinant represents the magnification produced by a linear transformation, we would expect the determinant of such a product of elementary matrices, representing the magnification produced by successive performance of several linear transformations, to equal the product of their separate magnifications. We shall show that this expectation is justified: the determinant of a product of matrices does equal the product of their determinants; and in doing so, we shall develop a method for evaluating determinants.

We first consider the effect of multiplying a matrix by one of the elementary matrices discussed in *Unit 3, Hermite Normal Form*. If we put an elementary matrix on the left of a given matrix and then multiply the matrices, the effect is the same as that of a row operation; but since we have defined determinants in terms of their columns, we want column operations rather than row operations. This can be achieved by placing the elementary matrices on the *right*, and then multiplying.

We illustrate this by examples (the classification of *column operations* and of elementary matrices is taken from pages N57 and N58). In each of the three cases, we suggest that you carry out the indicated matrix multiplication and fill in the entries in the blank matrix, to verify that the multiplication does have the same effect as the stated column operation.

Type I column operation (multiply a column by a scalar)

$$\begin{bmatrix} p & q & r \\ u & v & w \\ x & y & z \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix} = \begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$$

This result is true whether or not $c = 0$, but the second matrix factor is called an elementary matrix only if $c \neq 0$.

Type II column operation (add a multiple of one column to another column)

$$\begin{bmatrix} p & q & r \\ u & v & w \\ x & y & z \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c & 0 & 1 \end{bmatrix} = \begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$$

Type III column operation (interchange two columns)

$$\begin{bmatrix} p & q & r \\ u & v & w \\ x & y & z \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$$

The effects of these three types of column operation on the determinant of the matrix can be found from the definition of a determinant. For example, in the case of a Type I operation, Condition I in the definition of a determinant tells us that $D(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ is linear in each of its n variables (the columns of the matrix), and so if one of these columns is multiplied by

c, then so is the determinant. The corresponding results for operations of Types II and III are contained in the next reading passage.

READ Section III-2, pages K682–687 but note that there are sections of this reading passage which are not required in any detail. These sections are indicated in the notes below.

Notes

(i) **Theorem III-1**, *page* K682 This can be thought of as a special case of the rule we have just derived for Type I column operations, in which we take $c = 0$.
(ii) *line 7 and footnote*, *page* K682 Be careful to distinguish between a function and its image. In the footnote a permutation is a one-one function, whereas in the text, the permutation is the image of a one-one function. In fact, the function p has been applied to the subscripts of the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and the resulting rearrangement of the vectors is called a permutation.
(iii) *lines 2–4*, *page* K683 If you have the time, you might like to prove this missing result. Probably the easiest way to do it is by contradiction. (*Hint*: if you perform a sequence of interchanges on $\{1, \ldots, n\}$ and the result is again $\{1, \ldots, n\}$, can the number of interchanges be odd?)
(iv) *line 11*, *page* K683 The details of the rest of the argument leading to **Theorem III-3** are not important, although well worth studying if you have the time. The method has been used a number of times already; the only really new thing is its generality.
(v) **Theorem III-3** *page* K684 The form of the result is important, as is the observation in the paragraph which follows. The remainder of pages K684–6 to the end of the proof of **Theorem III-4** are not important, although it would be advisable to read them through, ignoring the technical details. We give an alternative proof of **Theorem III-4** in sub-section 5.1.5.
(vi) The remainder of Section III-2 beginning at "The last theorem . . ." on page K686 should be studied properly. (In spite of lines 4–6 on page K687, we shall not be studying the next section in K, but we do obtain the result in the next section of the correspondence text.)

We can summarize the effects of column operations by the following rules:

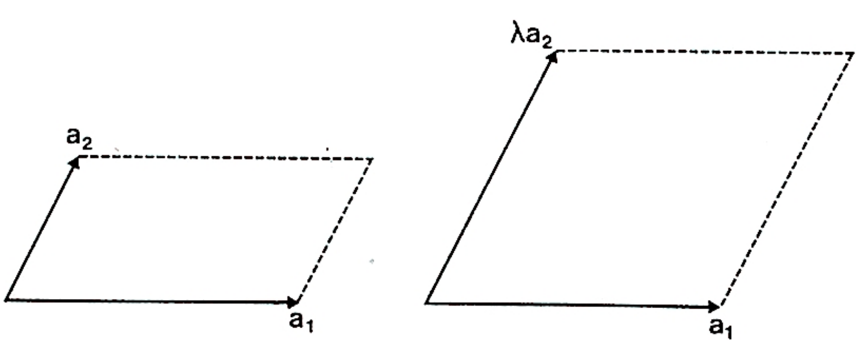
I Multiplying any column by a scalar multiplies the determinant by that scalar.

II Adding a multiple of any column to any other column leaves the determinant unaffected.

III Interchanging any two columns reverses the sign of the determinant.

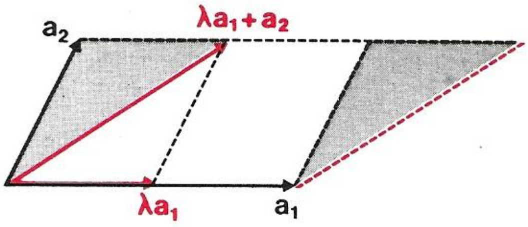
We can interpret these three rules geometrically in the case of endomorphisms of the plane, which we considered in sub-section 5.1.1.

I



If we multiply one side of a parallelogram by a factor, λ, then the area is multiplied by the same factor, λ. Here the columns of the determinant are represented by vectors $\mathbf{a}_1$ and $\mathbf{a}_2$.

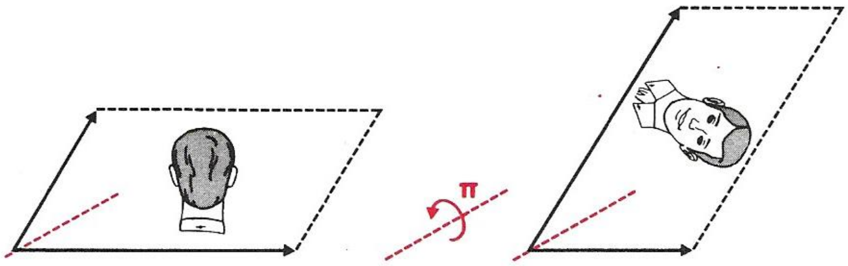
II



Since the shaded areas are equal, the parallelograms formed by a_1, a_2 and a_1, $\lambda a_1 + a_2$ have the same area

This corresponds to the elementary geometrical result that parallelograms on the same base and with equal heights have equal areas.

III



We can interchange $\mathbf{a}_1$ and $\mathbf{a}_2$ by reflecting in the line bisecting the angle between them. The result is to turn the parallelogram over (rotation of π), which accounts for the change of sign, as we have seen previously.

Using these rules we can evaluate any determinant. The following three examples show how.

Example 1

$$\begin{vmatrix} 4 & 3 \\ 2 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 3 \\ 0 & 1 \end{vmatrix} \quad \text{col. } 1 - 2 \times \text{col. } 2 \quad \text{(Rule II)}$$

$$= \begin{vmatrix} -2 & 0 \\ 0 & 1 \end{vmatrix} \quad \text{col. } 2 + \tfrac{3}{2} \times \text{col. } 1 \quad \text{(Rule II)}$$

$$= -2\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \quad \begin{array}{l}\text{using Rule I to remove the} \\ \text{factor } -2 \text{ from col. } 1\end{array}$$

$$= -2,$$

since $D(\mathbf{e}_1, \mathbf{e}_2) = 1$, by Condition III of the definition of the determinant function (page **K**680).

The above calculation could have been shortened if we had used the fact noted in Solution 3 of sub-section 5.1.2, that the determinant of an upper triangular matrix equals the product of its diagonal elements. This fact would have told us after the very first column operation that the determinant is -2. The standard method for evaluating determinants makes use of this fact.

Example 2

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -2 \\ 3 & 4 & 1 \\ 6 & 7 & 1 \end{vmatrix} \quad \text{col. 3 − col. 2 (the 1 at the bottom right simplifies the clearing of the bottom row)}$$

$$= \begin{vmatrix} 13 & 2 & -2 \\ -3 & 4 & 1 \\ 0 & 7 & 1 \end{vmatrix} \quad \text{col. 1} - 6 \times \text{col. 3}$$

$$= \begin{vmatrix} 13 & 16 & -2 \\ -3 & -3 & 1 \\ 0 & 0 & 1 \end{vmatrix} \quad \text{col. 2} - 7 \times \text{col. 3 (bottom row is now clear)}$$

$$= \begin{vmatrix} -3 & 16 & -2 \\ 0 & -3 & 1 \\ 0 & 0 & 1 \end{vmatrix} \quad \text{col. 1 − col. 2 (upper triangular)}$$

$$= (-3) \times (-3) \times 1 = 9.$$

Example 3

$$\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

(Two applications of Rule III: each column interchange reverses the sign.)

Exercises

Evaluate the following determinants by using column operations (if necessary) to reduce them to triangular form.

1. The determinant of the Type I elementary matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & c & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

2. The determinant of the Type II elementary matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ c & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

3. The determinant of the Type III elementary matrix:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

4. (i) $\begin{vmatrix} 1 & 3 & 0 \\ 0 & 1 & 2 \\ -1 & -2 & 1 \end{vmatrix}$ (ii) $\begin{vmatrix} 1 & 0 & -1 \\ 3 & 1 & -2 \\ 0 & 2 & 1 \end{vmatrix}$ (iii) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 0 & 0 \end{vmatrix}$

For Exercises 1, 2 and 3, compare your results with the factor by which the corresponding elementary operation multiplies the determinant of an arbitrary matrix, as given in Rules I, II and III above. Do you see any connection between Exercises 4(i) and 4(ii)?

Solutions

1. Determinant $= c$. It is already triangular.
2. Determinant $= 1$. It is already triangular.
3. Determinant $= -1$ (interchange cols. 1 and 4).
   In each case the determinant of the elementary matrix equals the corresponding factor given in Rules I, II, and III.
4. (i) Determinant $= -1$. Here is one of many methods:

$$\begin{vmatrix} 1 & 3 & 0 \\ 0 & 1 & 2 \\ -1 & -2 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 3 & 0 \\ 2 & 1 & 2 \\ 0 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 0 \\ 2 & 5 & 2 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} -\frac{1}{5} & 3 & 0 \\ 0 & 5 & 2 \\ 0 & 0 & 1 \end{vmatrix} = -1.$$

(col. 1 + col. 3) (col. 2 + 2 × col. 3) (col. 1 $-\frac{2}{5}$ × col. 2)

(ii) Determinant $= -1$.

The matrix whose determinant we evaluated in this exercise is the transpose of the one in Exercise 4 (i), and their determinants are equal.

(iii) Determinant $= -3$. One method is

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 0 & 0 \end{vmatrix} = 3\begin{vmatrix} 1 & 2 & 1 \\ 4 & 5 & 2 \\ 1 & 0 & 0 \end{vmatrix} = 3\begin{vmatrix} 1 & 0 & 1 \\ 4 & 1 & 2 \\ 1 & 0 & 0 \end{vmatrix}$$

(take 3 as common factor from col. 3) (col. 2 − 2 × col. 3)

$$= -3\begin{vmatrix} 1 & 0 & 1 \\ 2 & 1 & 4 \\ 0 & 0 & 1 \end{vmatrix} = -3\begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{vmatrix} = -3.$$

(interchange cols. 1 and 3) (col. 1 − 2 × col. 2)

5.1.4 Determinants and Matrix Multiplication

In the previous sub-section we saw how an elementary column operation multiplies the value of a determinant by a factor which depends only on the elementary operation (for example, interchanging two columns multiplies the determinant by the factor -1, regardless of the entries in the interchanged columns). Since any column operation is equivalent to multiplication on the right by some elementary matrix E, we can write this

$$\det(AE) = \det(A) \times k \qquad (1)$$

where k depends only on E.

We also saw (in Solutions 1, 2 and 3 of the previous sub-section) that for each typical elementary matrix the factor k is equal to the determinant of the elementary matrix. To see that this is true in general, we can take A in Equation (1) to be the unit matrix (since k does not depend on A) and use the fact that the determinant of the unit matrix is 1: then Equation (1) reduces to

$$\det(E) = k$$

and combining this result with (1) we get

$$\det(AE) = \det(A) \times \det(E). \qquad (2)$$

From this equation we can deduce two important theorems. The second shows det to be a morphism with respect to multiplication of matrices in its domain and multiplication of real numbers in its codomain, and the first justifies the introduction of determinants at this stage: it gives us a theoretical criterion telling us when a matrix or endomorphism is singular. We shall use this theorem later in this unit to give a method for calculating the eigenvalues of small (2×2 or perhaps 3×3) matrices. (It is equivalent to *Theorem 2.7* on page N91, but our proof is slightly different because we have defined the determinant function in a different way. It is also the same as **Theorem III-8** on page K689, which has a *completely* different proof.)

Theorem 1

det $A \neq 0$ if and only if A is non-singular.

Proof Notice the "if and only if". There are two things to prove:

(i) if A is non-singular, then det $A \neq 0$;
(ii) if det $A \neq 0$, then A is non-singular.

(i) If A is non-singular, it can be written as a product of elementary matrices (see *Theorem 6.1* on page N58), say $A = E_1 \cdots E_k$, and Equation (2) gives

$$\begin{aligned}\det(E_1E_2\cdots E_k) &= \det(E_1\cdots E_{k-1})\det(E_k)\\ &= \det(E_1\cdots E_{k-2})\det(E_{k-1})\det(E_k)\\ &= \det(E_1)\det(E_2)\cdots\det(E_k) \qquad (3)\end{aligned}$$

But we have seen that any elementary matrix can be obtained from the identity matrix by a column operation of Type I, II or III. The corresponding rules, as summarised on page 14, show that since det $I = 1$, the determinants of elementary matrices are never zero, and hence det $A \neq 0$.
(ii) We now have to show that if det $A \neq 0$, then A is non-singular. This is equivalent to the contrapositive proposition:

if A is singular, then det $A = 0$.

(See *Unit M100 17, Logic II*: the contrapositive of $\mathbf{a} \Rightarrow \mathbf{b}$ is $\sim \mathbf{b} \Rightarrow \sim \mathbf{a}$.) If A is singular, then its rank is less than n, the dimension of the domain vector space (see page N46, if necessary). Now the rank of A is equal to the maximum number of linearly independent columns (see page N41, if

necessary), so when A is singular, the columns of A must be linearly dependent. But we have already seen (**Theorem III-6**, page K686) that in such a case det $A = 0$. This completes the proof of the theorem.

Notice that an equivalent form of the statement of this theorem, as suggested in part (ii) above, is

det $A = 0$ if and only if A is singular.

Also notice that the first part of **Theorem III-8**, page K689, is an equivalent result: we have set it as an exercise.

The second theorem we shall deduce here is the multiplication rule for determinants, mentioned at the beginning of section 5.1.3. It is the same as ***Theorem 2.8***, page N91, and the proof is the same too. In **K** it is **Theorem III-10**, page K698, with a different proof.

Theorem 2

If A and B are $n \times n$ matrices, then

$$\det(AB) = \det(BA) = \det(A)\det(B).$$

Proof If A and B are non-singular, then each can be expressed as a product of elementary matrices, and the theorem follows by Equation (3) above; for example, if $A = E_1E_2$ and $B = E_3E_4$, we have

$$\det(AB) = \det(E_1E_2E_3E_4) = \det(E_1)\det(E_2)\det(E_3)\det(E_4)$$

and

$$\det(A)\det(B) = \det(E_1E_2)\det(E_3E_4) = \text{above expression,}$$

etc.

If either A or B is singular, then so are AB and BA (see, for example, ***Theorem 2.1*** on page N41), and hence by the preceding theorem det (AB), det (BA) and det (A) det (B) are all zero.

Exercises

1. Is it true that the determinant function is a morphism with respect to the operations of matrix multiplication in its domain and ordinary multiplication in its codomain?
2. Same as Exercise (1) but with "addition" for "multiplication".
3. Is it true that $\det(A^{-1}) = (\det(A))^{-1}$ for all non-singular square matrices A? Note that the -1's mean different things on the two sides of the equation.
4. Same as Exercise (3) but with "singular" for "non-singular".
5. Is it true that similar matrices always have the same determinant? (Similarity is defined on page N52.)
6. Is it true that if σ is an endomorphism of a finite-dimensional vector space V, represented by matrices A and A' with respect to two different bases in V, then $\det(A) = \det(A')$?
7. How would you define the determinant of an endomorphism σ of a finite-dimensional vector space V? (Make sure your definition is basis-independent.)
8. Prove the statement in the first sentence of **Theorem III-8**, page K689.

Solutions

1. Yes. This is just another way of stating our Theorem 2.

2. No. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

but, $\det \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \neq \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

3. Yes. Theorem 2 gives

$$\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det I = 1.$$

4. No. For a singular matrix, A, A^{-1} does not exist, and $(\det(A))^{-1}$ is meaningless because $\det A = 0$.

5. Yes. A' and A are similar if $A' = P^{-1}AP$ for some non-singular matrix P, and Theorem 2 gives $\det(P^{-1}AP) = \det(P^{-1}) \det(A) \det(P) = \det(A)$, since $\det(P^{-1}) \det(P) = 1$, by Solution 3.

6. Yes. If P is the matrix of transition from the basis for which the matrix representation of σ is A to the basis for which it is A', then we have $A' = P^{-1}AP$ by the rule given in the last paragraph on page N52. Hence, $\det(A') = \det(A)$ by Solution 5.

7. The determinant of σ can be defined as $\det A$ where A is the matrix representing σ with respect to an arbitrary basis in V; this is basis-independent by Solution 6.

8. We have to prove that "*A necessary and sufficient condition* (if and only if) *that a set of n vectors* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *in* $\mathscr{R}^n$ *be linearly dependent is that* $D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$.

   (i) If the vectors are linearly dependent, **Theorem III-6**, page K686, proves that $D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$.

   (ii) If $D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$, that is $\det A = 0$, for the corresponding matrix A, then our Theorem 1 states that A is singular; i.e. the rank of A is less than n. Thus the columns of A are linearly dependent.

5.1.5 The Determinant of the Transpose

In Exercises 4 (i) and (ii) of sub-section 5.1.3 we saw an example of a matrix with the same determinant as its transpose. As a third and final consequence of our result

$$\det(AE) = \det(A)\det(E)$$

we can deduce that *every* matrix* has the same determinant as its transpose:

$$\det(A) = \det(A^T) \qquad (1)$$

If A is singular, then so is its transpose (since row and column rank are equal), and so both sides of Equation (1) are 0. If A is non-singular, then it can be written as a product of elementary matrices

$$A = E_1 E_2 \cdots E_k$$

so that

$$A^T = E_k^T \cdots E_2^T E_1^T,$$

using the result that $(AB)^T = B^T A^T$ (see Exercise 4, page N57). Elementary matrices of Types I and III are equal to their transposes. Elementary matrices of Type II are upper or lower triangular and have determinant 1; their transposes also have determinant 1. Therefore we have

$$\begin{aligned}\det(A) &= \det(E_1)\det(E_2)\cdots\det(E_k)\\ &= \det(E_1^T)\det(E_2^T)\cdots\det(E_k^T)\\ &= \det(A^T).\end{aligned}$$

This result has the corollary that row operations on a matrix affect the determinant in the same way as the corresponding column operations. In practice we mix row and column operations to suit our convenience.

Exercises

1. For a general $n \times n$ matrix A, express $\det(-A)$ in terms of $\det(A)$.
2. A square matrix is said to be skew-symmetric if $A = -A^T$. What can be said about the determinant of skew-symmetric $n \times n$ matrices when n is odd?

Solutions

1. To convert A into $-A$, we need n column operations, each of which multiplies the determinant by -1; therefore $\det(-A) = (-1)^n \det(A)$. For example,

$$\begin{vmatrix} -1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{vmatrix} = +\begin{vmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{vmatrix}$$
$$= -\begin{vmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{vmatrix}$$

2. $\det(A) = \det(-A^T)$

$$\begin{aligned} &= \det(-A) && \text{by Equation 1 above}\\ &= (-1)^n\det(A) && \text{by Solution 1}\\ &= -\det(A) && \text{since } n \text{ is odd.}\end{aligned}$$

It follows that $\det(A) = 0$; i.e. any square skew-symmetric matrix of odd order is singular. (If the entries in the matrix

* Remember that determinants are defined for square matrices only.

are drawn from fields other than R, this result may not hold. For example, $\{0, 1\}$ with the usual arithmetic modulo 2 is a field for which $1 + 1 = 0$. Hence, from

$$\det A + \det A = 0$$

we cannot deduce that $\det A = 0$.)

5.1.6 Summary of Section 5.1

In this section we defined the terms

endomorphism	(page C5)	* * *
determinant function	(page K680)	* *
determinant of a $n \times n$ matrix	(page K681)	* * *
permutation	(page K683)	*
column operations	(page N57)	* * *

Theorems

1. **(III-1,** page K682)
   If one of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is the zero vector, then $D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$. * * *

2. **(III-2,** page K682)
   If the sequence of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is obtained from $\mathbf{a}_1, \ldots, \mathbf{a}_n$ by interchanging $\mathbf{a}_i$ and $\mathbf{a}_j$ $(i \neq j)$ then * * *
   $$D(\mathbf{b}_1, \ldots, \mathbf{b}_n) = -D(a_1, \ldots, \mathbf{a}_n).$$

3. **(Corollary,** page K682)
   If the sequence of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is obtained from $\mathbf{a}_1, \ldots, \mathbf{a}_n$ by shifting one of the $\mathbf{a}_i$ k places to the left or right, then *
   $$D(\mathbf{b}_1, \ldots, \mathbf{b}_n) = (-1)^k D(\mathbf{a}_1, \ldots, \mathbf{a}_n).$$

4. **(III-3,** page K684)
   For each n, there is one and only one function D satisfying Properties I through III of Definition III-1 (page K680). Its values are given by the formula *
   $$D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \sum_p \sigma(p) a_{p(1)1} a_{p(2)2} \cdots a_{p(n)n}.$$

5. **(III-4,** page K685)
   Let $M = (a_{ij})$ be an $n \times n$ matrix and let $M' = (b_{ij})$ be the transpose of M, i.e. the matrix whose columns are the rows of M (thus $b_{ij} = a_{ji}$). Then * * *
   $$\det M = \det M'.$$

6. **(III-5,** page K686)
   The value of a determinant is not changed by adding a multiple of the jth column (row) to the ith column (row) if $i \neq j$. * * *

7. **(III-6,** page K686)
   If the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly dependent, then * * *
   $$D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} = 0.$$

8. **(Theorem 1,** page C18)
   $\det A \neq 0$ if and only if A is non-singular. * * *

9. **(Theorem 2,** page C19)
   If A and B are $n \times n$ matrices, then * * *
   $$\det(AB) = \det(BA) = \det(A)\det(B).$$

Techniques

Evaluate the determinant of an $n \times n$ matrix by applying Rules I to III (page C14): * * *

I multiplying any column by a scalar multiplies the determinant by that scalar;

II adding a multiple of any column to any other column leaves the determinant unaffected;

III interchanging any two columns reverses the sign of the determinant,

and by using the fact that the determinant of an upper or lower triangular matrix equals the product of the entries on its main diagonal.

Notation

D	(page K680)
det M	(page K681)
σ	(page K683)

5.2 EIGENVALUES

5.2.1 Eigenvectors

We have already seen, in the Introduction to this unit, informal definitions of the terms eigenvalue and eigenvector. If an endomorphism L leaves the direction of a vector $\mathbf{x}$ in its domain unchanged, then we call $\mathbf{x}$ an *eigenvector* of L. In this case, L maps $\mathbf{x}$ to some scalar multiple of itself, say $\lambda\mathbf{x}$, and the value of the multiplier is called the corresponding *eigenvalue* of L. The algebraic version of this statement is the starting point of the next reading passage.

Example

Let $\begin{bmatrix} 1 & -4 \\ 0 & -1 \end{bmatrix}$ represent a linear transformation, σ, with respect to some fixed basis. Since

$$\begin{bmatrix} 1 & -4 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-1)\begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector of σ and -1 is the corresponding eigenvalue. You can check that, for any a, $\begin{bmatrix} 2a \\ a \end{bmatrix}$ is also an eigenvector corresponding to the eigenvalue -1.

Example

If D denotes the differentiation operator in $C^{\infty}(R)$, then

$$(D^2 - 1)\sin x = -2\sin x,$$

so that $x \longmapsto \sin x$ is an eigenvector of $D^2 - 1$ with -2 as the corresponding eigenvalue.

READ Section 12-2, page K461 *as far as line 4 of page* K462.

Notes

(i) *line −14, page* K461 In this unit we are looking at endomorphisms of vector spaces; so we shall take it that $\mathcal{S}$ and $\mathcal{V}$ are identical. The more general case considered in the book will not be required until *Unit 25, Boundary-value Problems*, in which the notion of eigenvalue is applied to differential equations.
(ii) *line −13, page* K461 "unknown parameter which may take real or complex values." All this means is that λ is a scalar, and that we are interested in the solution sets of Equation (12-9) for various values of λ. That is, we are interested in the sets $\{\mathbf{x}: L\mathbf{x} = \lambda\mathbf{x}\}$ for various λ.
(iii) *line −11, page* K461 "non-trivial solutions" means solutions other than the zero vector.
(iv) *line −9, page* K461 The remark about Euclidean spaces is not relevant to the present unit.

Exercises

1. In the vector space R^2, which of the following vectors are eigenvectors of the endomorphism whose matrix representation with respect to the standard basis is $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$? Where the given vector is an eigenvector, state also the corresponding eigenvalue.

   (a) $(1, 0)$, (b) $(0, 1)$, (c) $(1, 1)$, (d) $(-1, -1)$,
   (e) $(1, -1)$, (f) $(0, 0)$.

2. If D denotes the differentiation operator with domain and codomain $C^\infty(R)$, which of the following are eigenvectors of the differential operator $D + 1$? State the corresponding eigenvalue, when there is one.

(a) $x \longmapsto e^x$

(b) $x \longmapsto \sin x$

(c) $x \longmapsto 2e^{-x}$

(d) $x \longmapsto x$

3. If $\mathbf{x}$ is an eigenvector of an endomorphism L, with eigenvalue λ, and n is a positive integer, must it be true that $\mathbf{x}$ is an eigenvector of L^n? If so, give the corresponding eigenvalue; if not, give a counter-example. Interpret your result geometrically.

4. What relation exists between the solution set of $L\mathbf{x} = \lambda\mathbf{x}$ and the kernel of $L - \lambda$, which is defined by $(L - \lambda)\mathbf{x} = L\mathbf{x} - \lambda\mathbf{x}$?

Solutions

1. (a) and (b) are not eigenvectors;
(c) and (d) are eigenvectors, both with eigenvalue 2;
(e) is an eigenvector, with eigenvalue 0;
(f) is not an eigenvector (see line 1 of page K462).

2. (a) is an eigenvector, with eigenvalue 2;

$$(D + 1)e^x = 2e^x.$$

(c) is an eigenvector, with eigenvalue 0;
(b) and (d) are not eigenvectors.

3. Yes, $\mathbf{x}$ is an eigenvector of L^n with eigenvalue λ^n; for example,

$$L^2\mathbf{x} = L(L\mathbf{x}) = L(\lambda\mathbf{x}) = \lambda L(\mathbf{x}) = \lambda^2\mathbf{x},$$
$$L^3\mathbf{x} = L(L^2\mathbf{x}) = L(\lambda^2\mathbf{x}) = \lambda^2 L\mathbf{x} = \lambda^3\mathbf{x},$$

and so on. The geometrical interpretation is this: L multiplies $\mathbf{x}$ without affecting its direction, so if we perform L n times we multiply $\mathbf{x}$ by λ^n without changing its direction.

4. The solution set of $L\mathbf{x} = \lambda\mathbf{x}$ and the kernel of $L - \lambda$ are identical; for the kernel of $L - \lambda$ is defined as

$$\{\mathbf{x}: (L - \lambda)\mathbf{x} = 0\} = \{\mathbf{x}: L\mathbf{x} - \lambda\mathbf{x} = 0\}$$
$$= \{\mathbf{x}: L\mathbf{x} = \lambda\mathbf{x}\},$$

the solution set of $L\mathbf{x} = \lambda\mathbf{x}$.

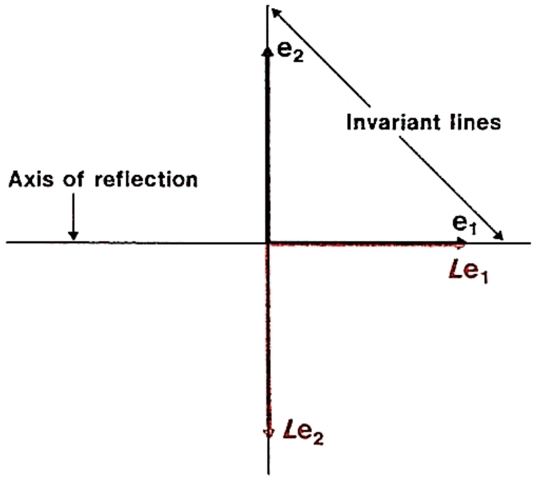
5.2.2 Eigenspaces and Invariant Subspaces

We have seen (Solution 4, sub-section 5.2.1) that the solution set of $L\mathbf{x} = \lambda\mathbf{x}$ is the kernel of $L - \lambda$. We know that the kernel is a subspace (page N31), and so the solution set of $L\mathbf{x} = \lambda\mathbf{x}$ is a subspace. The next reading passage exploits this latter fact to characterize eigenvalues and eigenvectors in a simple way.

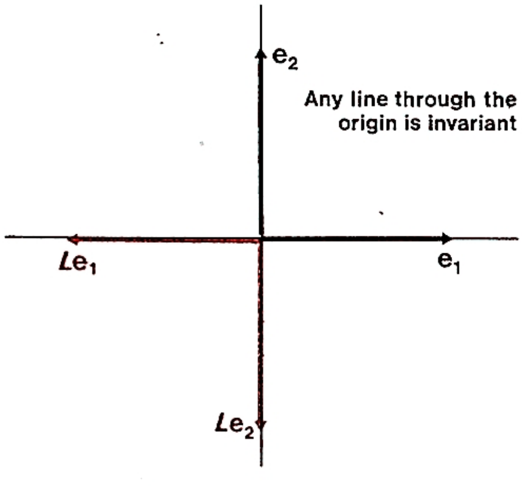
READ from page K462, *line 5, to the end of Example 3 on page* K463.

Notes

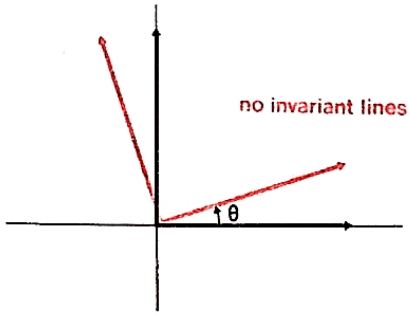
(i) *Lemma 12-1, page* K462 We know already that the solution set of $L\mathbf{x} = \lambda\mathbf{x}$ is a subspace for any scalar λ. The point of the lemma is that λ is an eigenvalue if and only if this subspace is non-trivial (i.e. contains non-zero vectors). $\mathcal{S}_{\lambda_0}$ is defined as the *eigenspace* of L corresponding to λ_0.
(ii) *line 12, page* K462 The zero vector is mentioned separately, because it is not allowed to be an eigenvector (line 1).
(iii) *line 14, page* K462 The case dim $\mathcal{S}_{\lambda_0} = 0$ is excluded since $\mathcal{S}_{\lambda_0}$ is required to be non-trivial.
(iv) **Definition 12-2,** *page* K462 In other words $\mathcal{W}$ is an invariant subspace if and only if $L(\mathcal{W}) \subseteq \mathcal{W}$.
(v) *line −8, page* K462 In other words, every eigenspace is an invariant subspace, but not every invariant subspace is an eigenspace. For example, $\{0\}$ is always an invariant subspace, but it is not an eigenspace.
(vi) *line −3, page* K462 The word "one-dimensional" is redundant in this definition since a space spanned by a single non-zero vector must be one-dimensional. Geometrically, this definition tells us that L maps the line $\langle\mathbf{x}\rangle$ to itself. If the eigenvalue is not 0, this mapping is onto; if it is 0, then $\langle\mathbf{x}\rangle$ maps to the zero vector; but in either case $\langle\mathbf{x}\rangle$ is invariant.
(vii) *Example 1, page* K463 Reflection in $\langle\mathbf{e}_1\rangle$



(viii) *Example 2, page* K463 Reflection across the origin (equivalent to 180° rotation in the plane)



(ix) *Example 3, page* K463 Rotation about the origin through an angle θ



Exercise

For a rotation about an axis through the origin in three-dimensional space, describe geometrically

(i) the invariant subspaces (and give their dimensions);
(ii) the eigenspaces (and give the corresponding eigenvalues).

Solution

(i) Assuming that the rotation is a multiple of π, the invariant subspaces are:

(a) the origin (dimension 0);
(b) the axis of rotation (dimension 1);
(c) the plane of rotation through the origin perpendicular to the axis of rotation (dimension 2);
(d) the entire space (dimension 3).

In addition, if the rotation is an odd multiple of π, then any line lying in the plane of rotation through the origin is an invariant subspace (dimension 1), as is any plane containing the axis of rotation (dimension 2).

If the rotation is an even multiple of π, then it is the identity and every subspace is invariant.

(ii) If the rotation is not a multiple of π, the only eigenspace is the axis of rotation, and its eigenvalue is 1. But if the rotation is an odd multiple of π, then also, as in Example 2, page K463, the plane of rotation through the origin is an eigenspace with eigenvalue -1.

If the rotation is an even multiple of π, then the eigenvalue is 1 and the eigenspace is the whole space.

5.2.3 Linear Independence of Eigenvectors

One of the main objects of studying eigenvectors of an endomorphism is that they can be used in constructing a basis which is particularly well adapted to that particular endomorphism. Since the one-dimensional invariant spaces give a set of "directions", it is natural to use these as far as possible as "axes" in setting up a coordinate system—or, in other words, to use the eigenvectors as at least some of the basis vectors. The first step in setting up such a basis is to make sure that the eigenvectors are linearly independent. The next reading passage gives a theorem that helps.

READ page K463 *from* "All this is simple enough . . ." *to the end of the section on page* K464.

Note

line -5, *page* K463 The method of proof by induction was discussed in *Unit M100 17, Logic II.*

Example

Two eigenvectors of a transformation represented by $\begin{bmatrix} 1 & -4 \\ 0 & -1 \end{bmatrix}$ are $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ corresponding to -1 and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ corresponding to $+1$. **Theorem 12-1** claims that these eigenvectors are linearly independent, as you can easily check.

Example

Notice that **Theorem 12-1** says nothing about eigenvectors corresponding to the same eigenvalue. These may or may not be linearly independent. Thus both

$$x \longmapsto \sin x \qquad (x \in R)$$

and

$$x \longmapsto \cos x \qquad (x \in R)$$

are eigenvectors of the linear differentiation operator $D^2 - 1$, corresponding to an eigenvalue -2. They are, however, linearly independent. (See also Exercise 2, below.)

Exercises

1. What is the largest possible number of distinct eigenvalues for an endomorphism of an n-dimensional vector space?
2. In **Theorem 12-1** on page K463, the eigenvalues are required to be distinct. Give a counter-example showing that the theorem would be false if the word "distinct" were omitted.

Solutions

1. Choose one eigenvector for each distinct eigenvalue. These eigenvectors are linearly independent; but the largest possible number of vectors in a linearly independent set is equal to the dimension of the space. So we can have at most n distinct eigenvalues.
2. Here are two possibilities:
   (a) Let $\mathbf{x}$ be an eigenvector of L and let $\mathbf{y}$ be a non-zero scalar multiple of $\mathbf{x}$: then $\mathbf{x}$ and $\mathbf{y}$ are eigenvectors of L, but are not linearly independent;
   (b) let L be the identity transformation in a vector space V; then any non-zero vector in V is an eigenvector of L, and any set of non-zero vectors in V is a set of eigenvectors of L, but that set is not necessarily linearly independent.

5.2.4 Eigenvector Bases

The theorem in the preceding sub-section gives a *partial* answer to the question: when can we form a basis consisting entirely of eigenvectors? If the vector space is n-dimensional, and the transformation has n different eigenvalues, then any set of corresponding eigenvectors, being linearly independent, constitutes a basis. The next reading passage gives the matrix representation of the endomorphism in such a basis, and gives an example of how it can be used for calculations.

READ Section 12–3 page K465 *to* "...has been left to the reader." *in the middle of page* K466.

Notes

(i) *Diagonal matrix, page* K465 If the basis is $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ and $\mathbf{a}_i$ is an eigenvector with eigenvalue λ_i, we have

$$L\mathbf{a}_1 = \lambda_1 \mathbf{a}_1, \text{ so that the first column is } \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$L\mathbf{a}_2 = \lambda_2 \mathbf{a}_2, \text{ so that the second column is } \begin{pmatrix} 0 \\ \lambda_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and so on. Remember that the coordinates of the image vectors of the basis are the columns of the matrix representation of L with respect to that basis.
(ii) **Theorem 12-2,** *page* K465 The condition that all eigenvalues be distinct is sufficient, but not necessary, for an eigenvector basis to exist. That is, an eigenvector basis can exist even when the number of distinct eigenvalues is less than n. For example, the identity transformation in $\mathcal{V}$ has the single eigenvalue 1, so that the theorem does not apply (unless $n = 1$). Since every non-zero vector is an eigenvector of the identity transformation, any basis of $\mathcal{V}$ is an eigenvector basis, and the identity transformation always has the same matrix representation

$$\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$

(iii) *Example 1, page* K465. Despite this example, the method is rarely used for linear problems in finite-dimensional spaces, because the methods based on row operations of matrices (see *Unit 3, Hermite Normal Form*) are more effective. For infinite-dimensional spaces, however, the method described in this example is often the only way to solve the problem.
(iv) *lines 9–14, page* K466 The case where the λ_i are all different from zero is the case where L is non-singular (see the second sentence on page K462). When L is singular, a solution is possible only if $\mathbf{y}$ lies in the image space of L, which is $\langle \mathbf{e}_2, \mathbf{e}_3 \rangle$ in the case considered here with only $\lambda_1 = 0$ (since

$$L(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3) = x_2\lambda_2\mathbf{e}_2 + x_3\lambda_3\mathbf{e}_3 \in \langle \mathbf{e}_2, \mathbf{e}_3 \rangle)$$

If $\mathbf{y}$ does lie in the image space, the general solution has the form:

{kernel} + particular solution.

In K's solution the term $x_1\mathbf{e}_1$ is a general element of the kernel $\langle \mathbf{e}_1 \rangle$, and the other two terms are a particular solution of $L\mathbf{x} = \mathbf{y}$.

Exercises

1. Under the conditions of **Theorem 12–2**, what is the matrix representing L^2 with respect to the eigenvector basis? If L is non-singular, what is the matrix representing L^{-1}, and how is it related to Example 1, page K465?

2. We saw, in Exercise 1 of sub-section 5.2.1, that the endomorphism L of R^2 whose matrix representation with respect to the standard basis is $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ has eigenvectors $(1, 1)$ and $(1, -1)$ corresponding to the eigenvalues 2 and 0 respectively. Write down the matrix of transition for changing to a basis consisting of these eigenvectors of L. Check your result!

3. Show that if L has matrix representation A with respect to some basis, and the space has a basis of eigenvectors, then $\det A$ is equal to the product of its eigenvalues. Interpret this result geometrically.

Solutions

1. L^2 is represented by $\begin{bmatrix} \lambda_1^2 & & & 0 \\ & \lambda_2^2 & & \\ & & \ddots & \\ 0 & & & \lambda_n^2 \end{bmatrix}$

L^{-1} is represented by $\begin{bmatrix} 1/\lambda_1 & & & 0 \\ & 1/\lambda_2 & & \\ & & \ddots & \\ 0 & & & 1/\lambda_n \end{bmatrix}$

In Example 1, the solution to $L\mathbf{x} = \mathbf{y}$ is $\mathbf{x} = L^{-1}\mathbf{y}$.

2. By formula (4.1) on page N50, the matrix of transition has columns giving the old coordinates of the new basis vectors; in the present case this matrix is

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \text{ or } \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

The formula for this change of basis is (see page N52, if necessary)

$$A' = P^{-1}AP.$$

To evaluate it, we use

$$AP = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}$$

and $P^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

Thus, the matrix in the new (eigenvector) basis is

$$A' = P^{-1}AP = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

which is in the form predicted by **Theorem 12–2.**

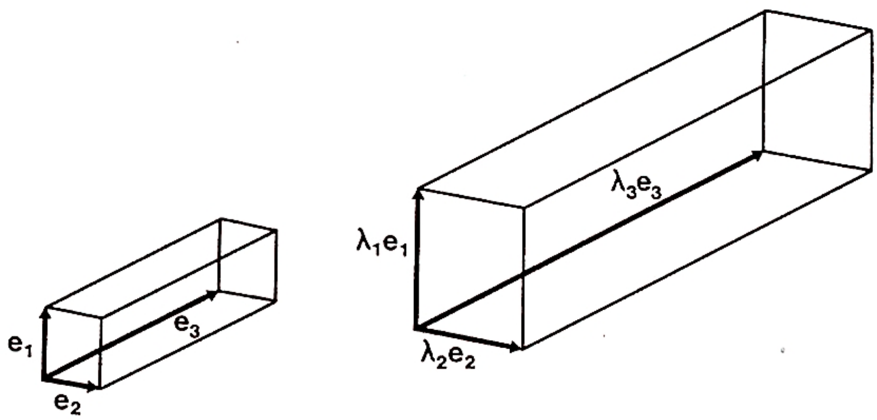
3. If B is the matrix representation of L with respect to the basis of eigenvectors, then

$$\det B = \det A$$

since the value of the determinant of a linear transformation is basis-independent (Solution 6, sub-section 5.1.4). The matrix B is diagonal and its entries are the eigenvalues; hence, $\det B$ equals the product of the eigenvalues, since B is diagonal (a special case of triangular).

The geometrical interpretation is this: the effect of L on an n-dimensional "parallelogram" with edges parallel to the n eigenvectors is to stretch the first edge by a factor λ_1, the second by a

factor λ_2, and so on, so that the volume of the "parallelogram" is multiplied by the product of these linear factors.



5.2.5 Summary of Section 5.2

In this section we defined the terms

non-trivial	(page C24)	* * *
eigenvector	(page K461)	* * *
eigenvalue	(page K461)	* * *
eigenspace	(page C26)	* * *
invariant under L	(page K462)	* * *
diagonal form of a matrix	(page K465)	* * *

Theorems

1. **(Lemma 12-1,** page K462)
The solution set of the equation $L\mathbf{x} = \lambda_0 \mathbf{x}$ is a nontrivial subspace of $\mathcal{S}$ for each eigenvalue λ_0 of L. ($L : \mathcal{S} \longrightarrow \mathcal{V}$) * * *

2. **(12-1,** page K463)
Any set of eigenvectors belonging to **distinct** eigenvalues for a linear transformation $L : \mathcal{S} \longrightarrow \mathcal{V}$ is linearly independent in $\mathcal{S}$. * * *

3. **(12-2,** page K465)
A linear transformation L mapping an n-dimensional vector space $\mathcal{V}$ into itself has at most n distinct eigenvalues. Moreover, when the number of distinct eigenvalues is equal to n, any complete set of eigenvectors, one for each eigenvalue, is a **basis** for $\mathcal{V}$, and the matrix of L with respect to such a basis is * * *

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

with the eigenvalues on the main diagonal and zeros elsewhere.

In the first and second theorems, we are interested in endomorphisms, i.e. L: $\mathcal{S} \longrightarrow \mathcal{S}$

Notation

$\mathcal{S}_{\lambda_0}$ (page K462)

5.3 SOME CALCULATIONS WITH EIGENVALUES

5.3.1 The Characteristic Polynomial

This section brings together the work in the first two sections. We shall use the determinant function to calculate eigenvalues.

READ from "The technique introduced in the above..." *on page* K466 *to the end of page* K469.

Notes

(i) *line 4, page* K467 "as we have already observed." This was in a section we have not asked you to read (page K459); but essentially the same observation was made in Solution 4 of sub-section 5.2.1.
(ii) *line 8, page* K467 Remember that I is represented by

$$\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}.$$

(iii) *Equation (12-14), page* K467 Another way of stating the argument is this: the scalar λ is an eigenvalue of L if and only if the linear transformation $L - \lambda$ is singular, and $L - \lambda$ is singular if and only if its determinant is zero. Equation (12-14) is just the expression for the determinant of $L - \lambda$.
(iv) *line −7, page* K467 The result that the determinant is a polynomial of degree n depends on **Theorem III-3**, page K684: there is just one product of the form $(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$ among the $n!$ products to be summed. The fact that the degree of the polynomial is n is very important in practice. It means that this method will work very well if $n = 2$. Since then the *characteristic equation* (12-14) is a quadratic; but for $n \geqslant 3$, the equation is at best a cubic and would normally have to be solved numerically. In fact, determinants larger than 3×3 are not usually suitable for numerical work, and so the eigenvalues of transformations represented by large matrices have to be found by other methods. We shall be looking at some of these methods later in the course.
(v) *bottom line, page* K468 This example has been contrived so that the characteristic equation can be solved easily, even though it is a cubic. One way of evaluating the determinant is this:

$$\begin{vmatrix} -\lambda & 0 & 1 \\ 0 & -(1+\lambda) & 0 \\ 2 & 2 & 1-\lambda \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ 0 & -(1+\lambda) & 0 \\ 2+\lambda-\lambda^2 & 2 & 1-\lambda \end{vmatrix}$$

col. $1 + \lambda \times$ col. 3

$$= - \begin{vmatrix} 1 & 0 & 0 \\ 0 & -(1+\lambda) & 0 \\ 1-\lambda & 2 & 2+\lambda-\lambda^2 \end{vmatrix}$$

interchange cols. 1 and 3

$$= (1+\lambda)(2+\lambda-\lambda^2).$$

This is 0 if $\lambda + 1 = 0$, or $2 + \lambda - \lambda^2 = 0$; that is, if $\lambda = -1$, or if $\lambda = -1$ or 2. The eigenvalues are -1 and 2.
(vi) *line 7, page* K469 A systematic way of dealing with such systems of equations is to use Hermite normal form. The matrix equation is equivalent to

$$\begin{bmatrix} -2 & 0 & 1 \\ 0 & -3 & 0 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Putting the matrix into Hermite normal form, gives

$$\begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which tells us that

$$x_1 = \tfrac{1}{2}x_3$$
$$x_2 = 0.$$

The solution set therefore comprises all vectors of the form $(\frac{1}{2}x_3, 0, x_3)$, with arbitrary x_3; i.e. it is $\langle \frac{1}{2}\mathbf{e}_1 + \mathbf{e}_3 \rangle$, or $\langle \mathbf{e}_1 + 2\mathbf{e}_3 \rangle$ as stated in **K**.

(vii) *line −3, page* K469 Complex solutions of this characteristic equation are not eigenvalues, since an eigenvalue is (by definition) a scalar and we are working with a real vector space.

Exercise

Find the eigenvalues and eigenspaces of the linear transformation in R^2 whose representation, with respect to the standard basis, is

(a) $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

If you need extra practice, try Exercise 1 on page K470.

Solutions

(a) The characteristic polynomial is

$$\begin{vmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} = (3-\lambda)(2-\lambda).$$

The characteristic equation is $(\lambda - 3)(\lambda - 2) = 0$, whose roots are 3 and 2. The eigenvectors corresponding to $\lambda = 3$ are the solutions of

$$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

which can be written

$$\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

So $x_1 = x_2$, and the eigenvectors are

$$x_1\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and the eigenspace is $\left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle$.

The eigenvectors corresponding to $\lambda = 2$ are the solutions of

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

So $x_2 = 0$, and the eigenvectors are

$$x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and the eigenspace is $\left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle$.

(b) The characteristic polynomial is

$$\begin{vmatrix} 1-\lambda & 1 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 - \lambda - 1.$$

The characteristic equation is $\lambda^2 - \lambda - 1 = 0$. Its solutions, the eigenvalues, are

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \simeq 1.618$$

and $\lambda_2 = \dfrac{1-\sqrt{5}}{2} \simeq -0.618$.

Rather than carry around ugly expressions, we will use the symbols λ_1, λ_2 respectively.

The eigenvectors (x_1, x_2) corresponding to the eigenvalue λ_1 are given by

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

i.e. $$\begin{bmatrix} 1-\lambda_1 & 1 \\ 1 & -\lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

If we multiply the first row by λ_1 and add the result to the second row, we get

$$\begin{bmatrix} 1-\lambda_1 & 1 \\ 1+\lambda_1-\lambda_1^2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

But λ_1 satisfies the characteristic equation, so that $\lambda_1^2 - \lambda_1 - 1 = 0$, i.e.

$$\begin{bmatrix} 1-\lambda_1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus

$$x_1 = \frac{x_2}{\lambda_1 - 1}.$$

But $\lambda_1^2 - \lambda_1 - 1 = 0$

i.e. $\lambda_1(\lambda_1 - 1) = 1.$

So $x_1 = \lambda_1 x_2$.

The eigenvectors are $(\lambda_1 x_2, x_2)$ and the eigenspace is

$$\langle \lambda_1 \mathbf{e}_1 + \mathbf{e}_2 \rangle.$$

Similarly the eigenspace corresponding to λ_2 is $\langle \lambda_2 \mathbf{e}_1 + \mathbf{e}_2 \rangle$.

5.3.2 An Application of Eigenvectors

One of the useful properties of an eigenvector basis is that not only the transformation itself but also its powers have a very simple matrix representation with respect to this basis. As an illustration of how this fact can be used, we apply it to get a formula for the general element of the Fibonacci sequence

$$1,1,2,3,5,8, \ldots, F_k, \ldots$$

The defining rule of this sequence is that each element is the sum of its two immediate predecessors, and the first two terms are both 1:

i.e. $F_{k+1} = F_k + F_{k-1}$ $(k = 2,3, \ldots)$.

The recurrence formula is equivalent to

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}.$$

In other words, we are saying that this recurrence formula can be expressed in terms of a linear transformation L in R^2

$$(F_{k+1}, F_k) = L(F_k, F_{k-1}) \tag{1}$$

where the matrix representation of L is

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

with respect to the standard basis.

By repeated application of Equation (1) we get

$$\begin{aligned} (F_3, F_2) &= L(F_2, F_1) = L(1, 1) \\ (F_4, F_3) &= L(F_3, F_2) = L^2(F_2, F_1) = L^2(1, 1) \\ (F_5, F_4) &= L(F_4, F_3) = L^3(1, 1) \end{aligned}$$

and in general

$$(F_{k+1}, F_k) = L^{k-1}(1, 1).$$

To calculate $L^{k-1}(1, 1)$ we express $(1, 1)$ in terms of a basis consisting of eigenvectors of L. The solution to the last exercise showed that two eigenvectors are

$$\mathbf{a}_1 = (\lambda_1, 1) \quad \text{and} \quad \mathbf{a}_2 = (\lambda_2, 1)$$

where $\lambda_1 = \frac{1}{2}(1 + \sqrt{5}) \simeq 1.618$

and $\lambda_2 = \frac{1}{2}(1 - \sqrt{5}) \simeq -0.618$. In terms of these, the vector $(1, 1)$ is given by

$$(1, 1) = \alpha_1(\lambda_1, 1) + \alpha_2(\lambda_2, 1)$$

where α_1 and α_2 are numbers which can be found by solving the simultaneous equations

$$(1, 1) = (\alpha_1\lambda_1 + \alpha_2\lambda_2, \alpha_1 + \alpha_2)$$

i.e.

$$\left.\begin{aligned} \alpha_1\lambda_1 + \alpha_2\lambda_2 &= 1 \\ \alpha_1 + \alpha_2 &= 1 \end{aligned}\right\}$$

so that $\alpha_1 = \dfrac{1 - \lambda_2}{\lambda_1 - \lambda_2}$ and $\alpha_2 = \dfrac{1 - \lambda_1}{\lambda_2 - \lambda_1}$.

The effect of L, or any power of L, on $(1, 1)$ is easily calculated, and so we have

$$\begin{aligned} L(1, 1) &= \alpha_1 L(\lambda_1, 1) + \alpha_2 L(\lambda_2, 1) \\ &= \alpha_1\lambda_1(\lambda_1, 1) + \alpha_2\lambda_2(\lambda_2, 1) \end{aligned}$$

and similarly

$$L^{k-1}(1, 1) = \alpha_1 \lambda_1^{k-1}(\lambda_1, 1) + \alpha_2 \lambda_2^{k-1}(\lambda_2, 1).$$

Thus,

$$(F_{k+1}, F_k) = \alpha_1 \lambda_1^{k-1}(\lambda_1, 1) + \alpha_2 \lambda_2^{k-1}(\lambda_2, 1);$$

that is, F_k, the kth Fibonacci number, is given by

$$F_k = \alpha_1 \lambda_1^{k-1} + \alpha_2 \lambda_2^{k-1}.$$

A test of this formula is given in the following table:

k	$\alpha_1 \lambda_1^{k-1}$	$\alpha_2 \lambda_2^{k-1}$	F_k
1	0.7236	0.2764	1
2	1.1708	−0.1708	1
3	1.8944	0.1056	2
4	3.0652	−0.0652	3
5	4.9597	0.0403	5
6	8.0249	−0.0249	8
7	12.9846	0.0154	13
8	21.0095	−0.0095	21
9	33.9941	0.0059	34
10	55.0037	−0.0036	55

One of the remarkable features of this formula is that for large k the first term alone gives a good approximation. This is a general feature of calculations involving high powers of matrices $\left(\text{here, high powers of} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right)$. As the power increases, the behaviour is dominated by the largest eigenvalue. This fact is the basis of the simplest practical method of numerically calculating eigenvalues and eigenvectors for matrices larger than 2 × 2. We shall be studying this method in *Unit 30, Numerical Solution of Eigenvalue Problems*.

Exercise

A method of approximating to $\sqrt{2}$, known to the ancient Greeks, is to use the sequence of fractions

$$\frac{1}{1}, \frac{3}{2}, \frac{7}{5}, \ldots, \frac{\alpha_n}{\beta_n}, \ldots \text{ in which}$$

$$\left.\begin{aligned} \alpha_{n+1} &= \alpha_n + 2\beta_n \\ \beta_{n+1} &= \alpha_n + \beta_n \end{aligned}\right\} (n = 1,2,\ldots).$$

Obtain explicit formulas for α_n and β_n and hence show that

$$\lim_{n \text{ large}} \left(\frac{\alpha_n}{\beta_n}\right) = \sqrt{2}.$$

Solution

In matrix form the recurrence relation is

$$\begin{bmatrix} \alpha_{n+1} \\ \beta_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_n \\ \beta_n \end{bmatrix}$$

or

$$(\alpha_{n+1}, \beta_{n+1}) = L(\alpha_n, \beta_n).$$

The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 2 \\ 1 & 1-\lambda \end{vmatrix} = 0$$

or $(1-\lambda)^2 - 2 = 0$, i.e. $\lambda^2 - 2\lambda - 1 = 0$.

Its solutions are λ_1 and λ_2, where

$$\lambda_1 = 1 + \sqrt{2}$$
$$\lambda_2 = 1 - \sqrt{2}$$

The corresponding eigenvectors are obtained by solving

$$\begin{bmatrix} -\sqrt{2} & 2 \\ 1 & -\sqrt{2} \end{bmatrix}\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and

$$\begin{bmatrix} \sqrt{2} & 2 \\ 1 & \sqrt{2} \end{bmatrix}\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$(\sqrt{2}, 1)$ and $(-\sqrt{2}, 1)$ are eigenvectors corresponding to the eigenvalues λ_1 and λ_2 respectively.

In terms of these as basis, the vector $(\alpha_1, \beta_1) = (1, 1)$ is

$$(1, 1) = \gamma_1(\sqrt{2}, 1) + \gamma_2(-\sqrt{2}, 1)$$

where $\sqrt{2}\gamma_1 - \sqrt{2}\gamma_2 = 1$

and $\gamma_1 + \gamma_2 = 1$

i.e. $\gamma_1 = (\sqrt{2} + 1)/(2\sqrt{2})$ and $\gamma_2 = (\sqrt{2} - 1)/(2\sqrt{2})$.

Then we have

$$(\alpha_n, \beta_n) = L^{n-1}(1, 1) = \gamma_1\lambda_1^{n-1}(\sqrt{2}, 1) + \gamma_2\lambda_2^{n-1}(-\sqrt{2}, 1)$$

i.e.

$$\alpha_n = \gamma_1\lambda_1^{n-1}\sqrt{2} - \gamma_2\lambda_2^{n-1}\sqrt{2}$$
$$\beta_n = \gamma_1\lambda_1^{n-1} + \gamma_2\lambda_2^{n-1}.$$

To calculate $\lim\limits_{n \text{ large}} \dfrac{\alpha_n}{\beta_n}$, we use the above forms.

Hence

$$\frac{\alpha_n}{\beta_n} = \frac{\gamma_1\lambda_1^{n-1}\sqrt{2} - \gamma_2\lambda_2^{n-1}\sqrt{2}}{\gamma_1\lambda_1^{n-1} + \gamma_2\lambda_2^{n-1}}.$$

Now divide numerator and denominator by $\gamma_1\lambda_1^{n-1}$ (λ_1, remember, is greater than λ_2)

$$\frac{\alpha_n}{\beta_n} = \frac{\sqrt{2} - (\gamma_2/\gamma_1)(\lambda_2/\lambda_1)^{n-1}\sqrt{2}}{1 + (\gamma_2/\gamma_1)(\lambda_2/\lambda_1)^{n-1}}$$
$$= \sqrt{2}\left(\frac{1 - (\gamma_2/\gamma_1)(\lambda_2/\lambda_1)^{n-1}}{1 + (\gamma_2/\gamma_1)(\lambda_2/\lambda_1)^{n-1}}\right).$$

Since $0 < \left|\dfrac{\lambda_2}{\lambda_1}\right| < 1$, $\lim\limits_{n \text{ large}} \dfrac{\alpha_n}{\beta_n} = \sqrt{2}$.

5.3.3 Summary of Section 5.3

In this section we defined the terms

characteristic polynomial	(page K467)	* * *
characteristic equation	(page K467)	* * *

Technique

Use the characteristic equation of a small $n \times n$ matrix to determine the eigenvalues and the corresponding eigenvectors of the transformation represented by the matrix. * * *

5.4 SUMMARY OF THE UNIT

Definitions

The terms defined in this unit and page references to their definitions are given below.

endomorphism	(page C5)	* * *
determinant function	(page K680)	* *
determinant of a $n \times n$ matrix	(page K681)	* * *
permutation	(page K683)	*
column operations	(page N57)	* * *
non-trivial	(page C24)	* * *
eigenvector	(page K461)	* * *
eigenvalue	(page K461)	* * *
eigenspace	(page C26)	* * *
invariant under L	(page K462)	* * *
diagonal form of a matrix	(page K465)	* * *
characteristic polynomial	(page K467)	* * *
characteristic equation	(page K467)	* * *

Theorems

We list the important theorems discussed in this unit. Only 3 star theorems which are essential to this and later units have been included in this list. References to the statements of the theorems are also given.

1. **(III-1**, page **K682)**
If one of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is the zero vector, then $D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$. * * *

2. **(III-2**, page **K682)**
If the sequence of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is obtained from $\mathbf{a}_1, \ldots, \mathbf{a}_n$ by interchanging $\mathbf{a}_i$ and $\mathbf{a}_j (i \neq j)$ then * * *

$$D(\mathbf{b}_1, \ldots, \mathbf{b}_n) = -D(\mathbf{a}_1, \ldots, \mathbf{a}_n).$$

3. **(III-4**, page **K685)**
Let $M = (a_{ij})$ be an $n \times n$ matrix and let $M' = (b_{ij})$ be the transpose of M, i.e., the matrix whose columns are the rows of M (thus $b_{ij} = a_{ji}$). Then * * *

$$\det M = \det M'.$$

4. **(III-5**, page **K686)**
The value of a determinant is not changed by adding a multiple of the jth column (row) to the ith column (row) if $i \neq j$. * * *

5. **(III-6**, page **K686)**
If the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly dependent, then * * *

$$D(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = 0.$$

6. **(Theorem 1**, page **C18)**
$\det A \neq 0$ if and only if A is non-singular. * * *

7. **(Theorem 2**, page **C19)**
If A and B are $n \times n$ matrices, then * * *

$$\det (AB) = \det (BA) = \det (A) \det (B).$$

8. **(Lemma 12-1,** page **K**462)
The solution set of the equation $L\mathbf{x} = \lambda_0 \mathbf{x}$ is a nontrivial subspace of $\mathcal{S}$ for each eigenvalue λ_0 of $L(L : \mathcal{S} \longrightarrow \mathcal{V})$. * * *

9. **(12-1,** page **K**463)
Any set of eigenvectors belonging to **distinct** eigenvalues for a linear transformation $L : \mathcal{S} \longrightarrow \mathcal{V}$ is linearly independent in $\mathcal{S}$. * * *

10. **(12-2,** page **K**465)
A linear transformation L mapping an n-dimensional vector space $\mathcal{V}$ into itself has at most n distinct eigenvalues. Moreover, when the number of distinct eigenvalues is equal to n, any complete set of eigenvectors, one for each eigenvalue, is a **basis** for $\mathcal{V}$, and the matrix of L with respect to such a basis is * * *

$$\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix},$$

with the eigenvalues on the main diagonal and zeros elsewhere.

Techniques

1. Evaluate the determinant of an $n \times n$ matrix by applying Rules I to III (page **C**14): * * *
   I multiplying any column by a scalar multiplies the determinant by that scalar;
   II adding a multiple of any column to any other column leaves the determinant unaffected;
   III interchanging any two columns reverses the sign of the determinant,

   and by using the fact that the determinant of an upper or lower triangular matrix equals the product of the entries on its main diagonal.

2. Use the characteristic equation of a small $n \times n$ matrix to determine the eigenvalues and the corresponding eigenvectors of the transformation represented by the matrix. * * *

Notation

D	(page **K**680)
det M	(page **K**681)
σ	(page **K**683)
$\mathcal{S}_{\lambda_0}$	(page **K**462)

5.5 SELF-ASSESSMENT

Self-assessment Test

This Self-assessment Test is designed to help you test quickly your understanding of the unit. It can also be used, together with the summary of the unit for revision. The answers to these questions will be found on the next non-facing page. We suggest you complete the whole test before looking at the answers.

1. By considering column and row operations, say how the determinants of the following matrices are related:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & 3 \\ 0 & 0 & 3 \end{bmatrix} \quad B = \begin{bmatrix} -2 & -1 & 3 \\ 2 & 1 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$

$$C = \begin{bmatrix} -1 & -2 & -3 \\ 1 & 2 & -3 \\ 0 & 0 & 3 \end{bmatrix}$$

2. Evaluate the determinants of the following matrices:

(i) $$A = \begin{bmatrix} 2 & 1 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & -3 \end{bmatrix}$$ (ii) $$B = \begin{bmatrix} 4 & 1 & 1 \\ -1 & 4 & 3 \\ 2 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

(iii) $$D = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3. Find the eigenvalues and eigenspaces of the endomorphisms of R^2 represented by the following matrices:

(i) $\begin{bmatrix} 4 & 0 \\ 2 & -1 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$

4. Determine the characteristic polynomial of the following matrix:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & 4 & 5 \end{bmatrix}$$

5. Exercise 4, page K464.

Solutions to Self-assessment Test

1. The row operation

add row 1 to row 2

applied to each of A, B and C yields an upper triangular matrix in each case. A zero element occurs in each main diagonal. Therefore

$$\det A = \det B = \det C = 0.$$

2. (i) This is upper triangular in form. Hence

$$\det A = 2 \times 1 \times (-3) = -6.$$

(ii) Subtract $2 \times$ row 3 from row 1, giving

$$\det B = \begin{vmatrix} 0 & 0 & 0 \\ -1 & 4 & 3 \\ 2 & \frac{1}{2} & \frac{1}{2} \end{vmatrix} = 0.$$

(iii) interchange col. 1 and col. 2, and interchange col. 3 and col. 4. Hence

$$\det D = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1.$$

3. (i) The characteristic equation is

$$\begin{vmatrix} 4-\lambda & 0 \\ 2 & -1-\lambda \end{vmatrix} = 0$$

where λ is real.

So $(4-\lambda)(1+\lambda) = 0$,

Hence $\lambda = 4$ and $\lambda = -1$ are the eigenvalues. Corresponding to $\lambda = 4$, the eigenvectors $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ are given by

$$\begin{bmatrix} 4 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

i.e.

$$\begin{bmatrix} 0 & 0 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Hence $2x_1 = 5x_2$, and the eigenvectors have the form $x_2 \begin{bmatrix} \frac{5}{2} \\ 1 \end{bmatrix}$ where x_2 is arbitrary.

The eigenspace is $\left\langle \begin{bmatrix} \frac{5}{2} \\ 1 \end{bmatrix} \right\rangle$.

Corresponding to $\lambda = -1$, the eigenvectors are given by

$$\begin{bmatrix} 5 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Hence $x_1 = 0$ and x_2 is arbitrary.

The eigenspace is $\left\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle$.

(ii) The characteristic equation is

$$\begin{vmatrix} -\lambda & 1 \\ 2 & -\lambda \end{vmatrix} = 0$$

i.e. $\lambda^2 = 2$,

So $\lambda = \sqrt{2}$ and $\lambda = -\sqrt{2}$.

Corresponding to $\lambda = \sqrt{2}$, the eigenvectors are given by

$$\begin{bmatrix} -\sqrt{2} & 1 \\ 2 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Hence $x_2 = \sqrt{2}x_1$

and $\sqrt{2}x_2 = 2x_1$.

Thus $x_2 = \sqrt{2}x_1$.

The eigenvectors have the form $x_1 \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$, where x_1 is arbitrary and the eigenspace is $\left\langle \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix} \right\rangle$.

Corresponding to $\lambda = -\sqrt{2}$, by similar reasoning to that for $\lambda = \sqrt{2}$, the eigenspace is $\left\langle \begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix} \right\rangle$.

(iii) The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} = 0$$

i.e. $(1-\lambda)(2-\lambda) - 2 = 0$

or $\lambda^2 - 3\lambda = 0$.

Hence $\lambda = 0$ and $\lambda = 3$.

Corresponding to $\lambda = 0$, the eigenvectors are given by

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

i.e. $x_1 + x_2 = 0$.

Thus, the eigenspace is $\left\langle \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\rangle$.

Corresponding to $\lambda = 3$, the eigenvectors are given by

$$\begin{bmatrix} -2 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

i.e. $-2x_1 + x_2 = 0$.

Thus, the eigenspace is $\left\langle \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\rangle$.

4. (i) The characteristic polynomial is

$$\begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 3 & 4 & 5-\lambda \end{vmatrix} = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & 0 & 1 \\ 3 & 4+5\lambda-\lambda^2 & 5-\lambda \end{vmatrix}$$

(add $\lambda \times$ col. 3 to col. 2)

$$= \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3+4\lambda+5\lambda^2-\lambda^3 & 4+5\lambda-\lambda^2 & 5-\lambda \end{vmatrix}$$

(add $\lambda \times$ col. 2 to col. 1)

$$= (-1)^2 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 + 5\lambda - \lambda^2 & 5 - \lambda & 3 + 4\lambda + 5\lambda^2 - \lambda^3 \end{vmatrix}$$

(two column interchanges)

$$= 3 + 4\lambda + 5\lambda^2 - \lambda^3.$$

5. (a) Let $\mathbf{x}$ be *any* element of the kernel (null space) of L. Then, by definition of $\mathbf{x}$

$$L\mathbf{x} = \mathbf{0}.$$

Hence L(kernel of L) $\subseteq$ kernel of L.

So, by **Definition 12-2** (page 462), the kernel is invariant.

(b) Let $\mathbf{x}$ be any element of $L(\mathcal{V})$; i.e. $\mathbf{x} = L\mathbf{y}$, for some $\mathbf{y} \in \mathcal{V}$.

Then $L(\mathbf{x}) = L(L\mathbf{y}) = L^2\mathbf{y}$

$= L\mathbf{y}$, since $L^2 = L$,

$= \mathbf{x}$, by definition of $\mathbf{y}$.

Hence $L\mathbf{x} \in L(\mathcal{V})$, and $L(\mathcal{V})$ is invariant by **Definition 12-2**.

LINEAR MATHEMATICS

335 01094 6